EXACTNESS OF MARTINGALE APPROXIMATION AND THE CENTRAL LIMIT THEOREM

Dalibor Volný

Université de Rouen

Abstract. Let (X_i) be a Markov chain with kernel Q, f an L^2 function on its state space. If Q is a normal operator and $f = (I-Q)^{1/2}g$ (which is equivalent to the convergence of $\sum_{n=1}^\infty \frac{\sum_{k=0}^{n-1} Q^k f}{n^{3/2}}$ in L^2), by Derriennic and Lin [D-L] we have the central limit theorem. By [M-Wu] and [Wu-Wo] the CLT is implied by the convergence of $\sum_{n=1}^\infty \frac{\| \sum_{k=0}^{n-1} Q^k f\|_2}{n^{3/2}}$ and by $\| \sum_{k=0}^{n-1} Q^k f\|_2 = o(\sqrt{n}/\log^q n)$, $q > 1$. We shall show that if Q is not normal or if the conditions of Maxwell and Woodroofe or Wu and Woodroofe are weakened by $\sum_{n=1}^\infty c_n \frac{\| \sum_{k=0}^{n-1} Q^k f\|_2}{n^{3/2}} < \infty$ for some sequence $c_n \searrow 0$ or by $\| \sum_{k=0}^{n-1} Q^k f\|_2 = O(\sqrt{n}/\log n)$, the CLT need not hold.

1. Introduction. Let $(\Omega, \mathcal{A}, \mu)$ be a probability space with a bijective, bimeasurable and measure preserving transformation T. For a measurable function f on ω, $(f \circ T^i)_i$ is a (strictly) stationary process and reciprocally, any (strictly) stationary process can be represented in this way.

Billingsley and Ibragimov (cf. [B], [I]) have proved that if $(m \circ T^i)$ is a martingale difference sequence with $m \in L^2$ and μ is ergodic (i.e. all sets $A \in \mathcal{A}$ for which $A = T^{-1}A$ it is $\mu(A) = 0$ or $\mu(A) = 0$) then $\frac{1}{\|m\|_2 \sqrt{n}} S_n(m)$ converge in law to the standard normal law $\mathcal{N}(0,1)$. Since the publication of Gordin's contribution [G], a special attention has been given to proving limit theorems via approximations by martingales. An important part of such results concern Markov chains, cf. e.g. [G-L], [K-V], [W], [D-L], [Wu-W].

Let $(S, \mathcal{B}, \mu)$ be a probability space, (ξ_i) a homogeneous and ergodic Markov chain with state space S, transition operator Q, and stationary distribution μ. For a measurable function f on S, $(f(\xi_i))$ is then a stationary random process; we shall study the central limit theorem for

$$S_n(f) = \sum_{i=0}^{n-1} f(\xi_i)$$

where $f \in L_0^2(\mu)$, i.e. is square integrable and has zero mean. Gordin and Lifšic ([G-L]) showed that if f is a solution of the equation

$$f = g - Qg$$

1991 *Mathematics Subject Classification.* 60G10, 60G42, 28D05, 60F05.

Key words and phrases. martingale aproximation, martingale difference sequence, strictly stationary process, Markov chain, central limit theorem.

with $g \in L^2$ then a martingale approximation giving the CLT exists. The result was extended to reversible operators Q and f satisfying

$$f = (I - Q)^{1/2} g \tag{1}$$

with $g \in L^2$ by Kipnis and Varadhan in [K-V], then for normal operators Q and f satisfying (1) by Derriennick and Lin in [D-L]. As noticed by Gordin and Holzmann ([G-H]), (1) is equivalent to the convergence of

$$\sum_{n=1}^{\infty} \frac{\sum_{k=0}^{n-1} Q^k f}{n^{3/2}} \quad \text{in} \quad L^2. \tag{2}$$

Maxwell and Woodroofe have shown in [M-W] that if

$$\sum_{n=1}^{\infty} \frac{\sum_{k=0}^{n-1} \|Q^k f\|_2}{n^{3/2}} < \infty \tag{3}$$

(without any other assumptions on the Markov operator Q) then the CLT takes place.

Remark that any stationary process can be represented by a Markov chain (cf. [Wu-W]). The central limit theorem of Maxwell and Woodroofe can thus be expressed in the following way:

Let $(\Omega, \mathcal{A}, P, T)$ be a probability space with a bimeasurable and measure preserving bijective transformation $T : \Omega \to \Omega$, $\mathcal{F}_i$ an increasing filtration with $T^{-1}\mathcal{F}_i = \mathcal{F}_{i+1}$, f is a square integrable and zero mean function on Ω, $\mathcal{F}_0$-measurable. We denote

$$S_n(f) = \sum_{i=0}^{n-1} f \circ T^i.$$

(2) then becomes

$$\sum_{n=1}^{\infty} \frac{E(S_n(f) \,|\, \mathcal{F}_0)}{n^{3/2}} \tag{2'}$$

and (3) becomes

$$\sum_{n=1}^{\infty} \frac{\|E(S_n(f) \,|\, \mathcal{F}_0)\|_2}{n^{3/2}} < \infty. \tag{3'}$$

In [Wu-W], Wu and Woodroofe have shown that if

$$\|E(S_n(f) \,|\, \mathcal{F}_0)\|_2 = o\Big(\frac{\sqrt{n}}{n^q}\Big) \tag{4}$$

for some $q > 1$ then the CLT takes place.

In [Vo] and [Kl-Vo1], nonadapted versions of Maxwell-Woodroofe aproximations (3') and have been found.

In the present paper we will deal with exactness of the central limit theorems of Derriennic and Lin, Maxwell and Woodroofe, and of Wu and Woodroofe. First, we show that the central limit theorem of Derriennic and Lin cannot be extended to non-normal operators Q.

Theorem 1. *There exists a process $(f \circ T^i)$ such that the series*

$$\sum_{n=1}^{\infty} \frac{E(S_n(f)\,|\,\mathcal{F}_0)}{n^{3/2}} \tag{2'}$$

converges in L^2, but for two different subsequences (n'_k), (n''_k), the distributions of $S_{n'_k}/\sigma_{n'_k}$ and $S_{n''_k}/\sigma_{n''_k}$ converge to different limits.

Then we show that in the central limit theorems of Maxwell and Woodroofe and of Wu and Woodroofe, the rate of convergence of $\|E(S_n(f)|\mathcal{F}_0)\|_2$ towards 0 is practically optimal.

Theorem 2. *For any sequence of positive reals $c_n \to 0$ there exists a process $(f \circ T^i)$ such that*

$$\sum_{n=1}^{\infty} c_n \frac{\|E(S_n(f)|\mathcal{F}_0)\|_2}{n^{3/2}} < \infty \tag{5}$$

but for two different subsequences (n'_k), (n''_k), the distributions of $S_{n'_k}/\sigma_{n'_k}$ and $S_{n''_k}/\sigma_{n''_k}$ converge to different limits.

In [Pe-U], Peligrad and Utev have shown that under the same assumptions there exists an f such the sequence of $S_n(f)/\sqrt{n}$ is not stochastically bounded. (Remark that in the same paper the authors have proved that (3') implies also the weak invariance principle.)

Theorem 3. *There exists a process $(f \circ T^i)$ such that*

$$\|E(S_n(f)|\mathcal{F}_0)\|_2 = O\Big(\frac{\sqrt{n}}{\log n}\Big), \tag{6}$$

but for two different subsequences (n'_k), (n''_k), the distributions of $S_{n'_k}/\sigma_{n'_k}$ and $S_{n''_k}/\sigma_{n''_k}$ converge to different limits.

From the construction it follows that in Theorems 1-3, the variances σ_n^2 of $S_n(f)$ grow faster than linearly. It thus remains an open problem whether with a supplementary assumption $\sigma_n^2/n \to$ *const.* the CLT would hold. As shown by a couter example in [Kl-Vo2], this assumption is not sufficient for $q \le 1/2$, the only exponents to consider are thus $1/2 < q < 1$.

It also remains an open question whether the CLT would hold for $f \in L^{2+\delta}$ for some $\delta > 0$.

2. Proof.

We give one proof which will treat all three theorems.

In all of the text, log will denote the dyadic logarithm.

For $k = 1, 2, \dots$ let $n_k = 2^k$, e_k be random variables with

$$\|e_k\|_2 = a_k/k, \quad 0 \le a_k \le 1, \quad \sum^{\infty} a_k/k = \infty,$$

such that for each k, $U^i e_k$ are independent, and if $i \neq j$ then $U^i e_{k'}$ and $U^j e_{k''}$ are orthogonal. For $k' \neq k''$ the $e_{k'}$, $e_{k''}$ are not orthogonal but we suppose that for all $1 \leq k', k''$ it is $E(e_{k'} e_{k''}) \geq 0$ and

$$\Big\| \sum_{k=1}^{n} e_k \Big\|_2 \nearrow \infty \quad \text{as} \quad n \to \infty.$$

Let

$$f = \sum_{k=1}^{\infty} \frac{1}{n_k} \sum_{i=0}^{n_k-1} U^{-i} e_k.$$

We have $\|f\|_2 \leq \sum_{k=1}^{\infty} \|e_k\|_2 / \sqrt{n_k} < \infty$ due to the exponential growth of the n_ks.

For a positive integer N we have

$$S_N(f) = \sum_{k=1}^{\infty} \sum_{j=0}^{N-1} \sum_{i=0}^{n_k-1} \frac{1}{n_k} U^{j-i} e_k = S_N'(f) + S_N''(f)$$

where

$$S_N'(f) = S_N(f) - E(S_N(f)|\mathcal{F}_0) = \sum_{k=1}^{\infty} \sum_{j=0}^{N-1} \sum_{i=0}^{(j\wedge n_k)-1} \frac{1}{n_k} U^{j-i} e_k$$

$(j \wedge n_k = \min\{j, n_k\})$ and

$$S_N''(f) = E(S_N(f)|\mathcal{F}_0) = \sum_{k=1}^{\infty} \sum_{j=0}^{N-1} \sum_{i=j}^{n_k-1} \frac{1}{n_k} U^{j-i} e_k.$$

We will study the asymptotic behaviour of $S_N''(f) = E(S_N(f)|\mathcal{F}_0)$ and $S_N'(f) = S_N(f) - E(S_N(f)|\mathcal{F}_0)$ separately. In the first case we will show that the series

$$\sum_{n=1}^{\infty} \frac{E(S_n(f)\,|\,\mathcal{F}_0)}{n^{3/2}}$$

converges in L^2, for a suitable choice of a_k we shall have

$$\|E(S_n(f)|\mathcal{F}_0)\|_2 = O\Big(\frac{\sqrt{n}}{\log n}\Big)$$

and for any sequence of positive reals $c_n \to 0$ the a_k can be chosen so that

$$\sum_{n=1}^{\infty} c_n \frac{\|E(S_n(f)|\mathcal{F}_0)\|_2}{n^{3/2}} < \infty.$$

In the second case we will show that the assumption $\sum_{k=1}^{\infty} a_k/k = \infty$ allows us to define the e_k so that for two different subsequences (n_k'), (n_k''), the distributions of $S_{n_k'}/\sigma_{n_k'}$ and $S_{n_k''}/\sigma_{n_k''}$ converge to different limits.

Eventually we prove existence of a dynamical system on which the process can be defined.

1. Asymptotics of $S''_N(f) = E(S_N(f)|\mathcal{F}_0)$.

For $N \le n_k$ we have

$$\sum_{j=0}^{N-1} \sum_{i=j}^{n_k-1} U^{j-i} e_k = \sum_{i=0}^{n_k-N} N U^{-i} e_k + \sum_{i=n_k-N+1}^{n_k-1} (n_k - i) U^{-i} e_k$$

and for $N > n_k$ we have

$$\sum_{j=0}^{N-1} \sum_{i=j}^{n_k-1} U^{j-i} e_k = \sum_{j=0}^{n_k-1} (n_k - j) U^{-j} e_k,$$

hence

$$S''_N(f) = \sum_{k\ge1:n_k<N} \sum_{j=0}^{n_k-1} \frac{n_k-j}{n_k} U^{-j} e_k + \\ + \sum_{k\ge1:n_k\ge N} \Big[\sum_{j=0}^{n_k-N} \frac{N}{n_k} U^{-j} e_k + \sum_{j=n_k-N+1}^{n_k-1} \frac{n_k-j}{n_k} U^{-j} e_k \Big]. \tag{7}$$

We will prove that (2') is satisfied. For this, it is sufficient to show that

$$\sum_{N=1}^{\infty} \sum_{k\ge1:n_k<N} \sum_{j=0}^{n_k-1} \frac{1}{N^{3/2}} \frac{n_k-j}{n_k} U^{-j} e_k$$

and

$$\sum_{N=1}^{\infty} \sum_{k\ge1:n_k\ge N} \frac{1}{N^{3/2}} \Big[\sum_{j=0}^{n_k-N} \frac{N}{n_k} U^{-j} e_k + \sum_{j=n_k-N+1}^{n_k-1} \frac{n_k-j}{n_k} U^{-j} e_k \Big]$$

converge in L^2.

Recall that $n_k = 2^k$. For the first sum we have

$$\Big\| \sum_{N=1}^{\infty} \sum_{k\ge1:2^k<N} \sum_{j=0}^{2^k-1} \frac{1}{N^{3/2}} \frac{2^k-j}{2^k} U^{-j} e_k \Big\|_2^2 =$$

$$\Big\| \sum_{k=1}^{\infty} \Big(\sum_{N=2^k+1}^{\infty} \frac{1}{N^{3/2}} \Big) \sum_{j=0}^{2^k-1} \frac{2^k-j}{2^k} U^{-j} e_k \Big\|_2^2 \le$$

$$\Big\| c \sum_{k=1}^{\infty} \sum_{j=0}^{2^k-1} \frac{2^k-j}{2^{3k/2}} U^{-j} e_k \Big\|_2^2 \le$$

$$\Big\| c \sum_{j=0}^{\infty} \sum_{k>\log(j+1)} \frac{2^k-j}{2^{3k/2}} U^{-j} e_k \Big\|_2^2 \le$$

$$\sum_{j=0}^{\infty} \Big\| c \sum_{k>\log(j+1)} \frac{1}{2^{k/2}} U^{-j} e_k \Big\|_2^2 \le C \sum_{j=0}^{\infty} \Big(\frac{1}{\sqrt{(j+1)[1 \vee \log(j+1)]}} \Big)^2 < \infty$$

where $0 < c, C < \infty$. For the second sum we have

$$\Big\| \sum_{N=1}^{\infty} \sum_{k\geq 1: 2^k \geq N} \frac{1}{N^{3/2}} \sum_{j=0}^{2^k-N} \frac{N}{2^k} U^{-j} e_k \Big\|_2^2 \leq$$

$$\Big\| \sum_{N=1}^{\infty} \sum_{j=0}^{\infty} \sum_{k\geq \log(N+j)} \frac{1}{N^{1/2}2^k} U^{-j} e_k \Big\|_2^2 \leq \Big\| \sum_{j=0}^{\infty} \sum_{k\geq \log(j+1)} \Big(\sum_{N=1}^{2^k} \frac{1}{N^{1/2}} \Big) \frac{1}{2^k} U^{-j} e_k \Big\|_2^2 \leq$$

$$\sum_{j=0}^{\infty} \Big\| c \sum_{k\geq \log(j+1)} \frac{1}{2^{k/2}} U^{-j} e_k \Big\|_2^2 \leq C \sum_{j=0}^{\infty} \Big(\frac{1}{\sqrt{(j+1)}[1 \vee \log(j+1)]} \Big)^2 < \infty$$

and

$$\Big\| \sum_{N=1}^{\infty} \sum_{k\geq 1: 2^k \geq N} \frac{1}{N^{3/2}} \sum_{j=2^k-N+1}^{2^k-1} \frac{2^k - j}{2^k} U^{-j} e_k \Big\|_2^2 \leq$$

$$\sum_{j=0}^{\infty} \Big\| \sum_{k\geq \log(j+1)} \sum_{N=2^k-j+1}^{2^k} \frac{1}{N^{3/2}} \frac{2^k - j}{2^k} U^{-j} e_k \Big\|_2^2 \leq$$

$$\sum_{j=0}^{\infty} \Big\| \sum_{k\geq \log(j+1)} \Big(\sum_{N=2^k-j+1}^{2^k} \frac{1}{N^{1/2}} \Big) \frac{1}{2^k} U^{-j} e_k \Big\|_2^2 \leq$$

$$C \sum_{j=0}^{\infty} \Big(\frac{1}{\sqrt{(j+1)}[\log(j+1)]} \Big)^2 < \infty$$

where $0 < c, C < \infty$. This finishes the proof of (2').

We have

$$\begin{aligned} &(1/\sqrt{6})\|e_k\|_2 \sqrt{n_k} \leq \Big\| \sum_{j=0}^{n_k-1} \frac{n_k - j}{n_k} U^{-j} e_k \Big\|_2 \leq \|e_k\|_2 \sqrt{n_k}, \quad n_k < N, \\ &\Big\| \sum_{j=0}^{n_k-N} \frac{N}{n_k} U^{-j} e_k + \sum_{j=n_k-N+1}^{n_k-1} \frac{n_k - j}{n_k} U^{-j} e_k \Big\|_2 \leq \frac{N}{\sqrt{n_k}} \|e_k\|_2, \quad n_k \geq N. \end{aligned} \tag{8}$$

Recall that by $[x]$ we denote the integer part of x. Because $n_k = 2^k$ grow exponentially fast, there exists a constant $0 < c < \infty$ not depending on N such that

$$\sum_{k\geq 1: n_k \geq N} (N/\sqrt{n_k}) \|e_k\|_2 \leq c \|e_{[\log N]}\|_2, \qquad \sum_{k \geq 1: n_k < N} \|e_k\|_2 \sqrt{n_k} \leq c\sqrt{N} \|e_{[\log N]}\|_2.$$

Using (7) and (8) we deduce that for some constants $c', c'' > 0$ we have

$$c' N \|e_{[\log N]}\|_2^2 < \|E(S_N(f)|\mathcal{F}_0)\|_2^2 < c'' N \|e_{[\log N]}\|_2^2.$$

Because $\|e_k\|_2 = a_k/k$,

$$c' \frac{N a_{[\log N]}^2}{\log^2 N} < \|E(S_N(f)|\mathcal{F}_0)\|_2^2 < c'' \frac{N a_{[\log N]}^2}{\log^2 N}. \tag{9}$$

For $a_k \equiv 1$ we thus have

$$\|E(S_n(f)|\mathcal{F}_0)\|_2 = O\Big(\frac{\sqrt{n}}{\log n}\Big).$$

From (9) we deduce that the series $\sum_{n=1}^{\infty} n^{-3/2}\|E(S_n(f)|\mathcal{F}_0)\|_2$ converges if and only if $\sum_{n=1}^{\infty} a_{[\log n]}/(n[\log n])$ converges; because

$$\frac{a_k}{2k} \le \sum_{j=0}^{2^k-1} \frac{a_k}{k(2^k+j)} \le \frac{a_k}{k}$$

this is equivalent to the convergence of $\sum_{n=1}^{\infty} a_n/n$.

Let c_n be positive real numbers, $c_n \to 0$; we can choose the a_n so that $\sum_{n=1}^{\infty} a_n c_n/n < \infty$, that means

$$\sum_{n=1}^{\infty} c_n \frac{\|E(S_n(f)|\mathcal{F}_0)\|_2}{n^{3/2}} \approx \sum_{n=1}^{\infty} \frac{c_n a_{[\log n]}}{n[\log n]} < \infty$$

but $\sum_{n=1}^{\infty} a_n/n = \infty$, i.e. .

$$\sum_{n=1}^{\infty} \frac{\|E(S_n(f)|\mathcal{F}_0)\|_2}{n^{3/2}} = \infty.$$

2. Asymptotics of $S'_N(f) = S_N(f) - E(S_N(f)|\mathcal{F}_0)$.

Notice that in the preceding section, no hypothesis on dependence of the e_k was needed. Now, we shall suppose that the sequence of a_k is fixed and we choose the e_k so that for two different subsequences (n'_k), (n''_k), the distributions of $S_{n'_k}/\sigma_{n'_k}$ and $S_{n''_k}/\sigma_{n''_k}$ converge to different limits.

For $N \le n_k$ we have

$$\sum_{j=0}^{N-1} \sum_{i=0}^{(j\wedge n_k)-1} U^{j-i} e_k = \sum_{j=1}^{N-1} (N-j) U^j e_k \tag{10}$$

and for $N > n_k$ we have

$$\sum_{j=0}^{N-1} \sum_{i=0}^{(j\wedge n_k)-1} U^{j-i} e_k = \sum_{j=1}^{N-n_k} n_k U^j e_k + \sum_{j=N-n_k+1}^{N-1} (N-j) U^j e_k. \tag{11}$$

For all $k \ge 1$ we have $P l U^j e_k = 0$ if $j \ne l$, $P l U^l e_k = U^l e_k$. For $l \ge N$ and $l \le 0$ we thus have $P_l S_N(f) = 0$ and for $1 \le l \le N-1$ we, using (10) and (11), deduce

$$P_l S_N(f) = \sum U^l e_k + \sum \frac{N-l}{n_k} U^l e_k. \tag{12}$$

Recall that $[x]$ denotes the integer part of x. We have

$$S'_N(f) = \sum_{l=1}^{N-1} P_l S_N(f) = \sum_{l=1}^{[N(1-\epsilon)]} P_l S_N(f) + \sum_{l=[N(1-\epsilon)]+1}^{N} P_l S_N(f)$$

where

$$\sum_{l=1}^{[N(1-\epsilon)]} P_l S_N(f) =$$

$$\sum_{l=1}^{[N(1-\epsilon)]} U^l \sum_{k\geq 1: n_k \leq N-l} e_k + \sum_{l=1}^{[N(1-\epsilon)]} U^l \sum_{k\geq 1: n_k \geq N+1-l} \frac{N-l}{n_k} e_k =$$

$$= \sum_{l=1}^{[N(1-\epsilon)]} U^l \sum_{k\geq 1: n_k \leq \epsilon N} e_k + \sum_{l=1}^{[N(1-\epsilon)]} U^l \sum_{k\geq 1: \epsilon N < n_k \leq N-l} e_k +$$

$$+ \sum_{l=1}^{[N(1-\epsilon)]} U^l \sum_{k\geq 1: n_k \geq N+1-l} \frac{N-l}{n_k} e_k.$$

Because $n_k = 2^k$,

$$\Big\| \sum_{k\geq 1: n_k \geq N+1-l} \frac{N-l}{n_k} U^l e_k \Big\|_2 \leq 2\|e_{[\log(N-l)]}\|_2 \leq 2/\log(N-l); \tag{13}$$

$\epsilon N < n_k \leq N$ if and only if $\log N + \log \epsilon < k \leq \log N$. We thus deduce that for $\epsilon > 0$ fixed and $b(N) = \| \sum_{k=1}^{\log N} e_k \|_2 \nearrow \infty$,

$$\lim_{N\to\infty} \frac{1}{b(N)\sqrt{N}} \Big\| \sum_{\epsilon N < n_k \leq N} e_k \Big\|_2 = 0$$

and

$$\lim_{N\to\infty} \frac{1}{b(N)\sqrt{N}} \Big\| \sum_{l=1}^{[N(1-\epsilon)]} P_l S_N(f) - \sum_{l=1}^{[N(1-\epsilon)]} U^l \sum_{k\geq 1: n_k \leq \epsilon N} e_k \Big\|_2 = 0.$$

For all $[N(1-\epsilon)] + 1 \leq l \leq N-1$ we have, by (12) and (13), $\|P_l S_N(f)\|_2 \leq b(N-l) + 2/\log(N-l)$ hence

$$\lim_{\epsilon \searrow 0} \lim_{N\to\infty} \frac{1}{b(N)\sqrt{N}} \Big\| \sum_{l=[N(1-\epsilon)]+1}^{N-1} P_l S_N(f) - \sum_{l=[N(1-\epsilon)]+1}^{N-1} \sum_{k\geq 1: n_k \leq \epsilon N} U^l e_k \Big\|_2 = 0,$$

therefore

$$\lim_{N\to\infty} \frac{1}{b(N)\sqrt{N}} \Big\| S'_N(f) - \sum_{l=0}^{N-1} U^l \sum_{k\geq 1: n_k \leq N} e_k \Big\|_2 = 0. \tag{14}$$

Recall that

$$\|e_k\|_2 = a_k/k, \quad 0 \leq a_k \leq 1, \quad \sum^{\infty} a_k/k = \infty.$$

Let N_l, $l = 1, 2, \dots$, be an increasing sequence of positive integers such that

$$2^{2^l} - 1 < \sum_{k\geq 1: N_{l-1} < n_k \leq N_l} \frac{a_k}{k} < 2^{2^l} + 1; \tag{15}$$

we suppose that for $N_{l-1} < n_k \leq N_l$ the random variables e_k are multiples one of another and are independent of any e_j with $n_j \leq N_{l-1}$ or $n_j > N_l$. For l odd we choose e_k, $N_{l-1} < n_k \leq N_l$, so that

$$\frac{1}{b(N_l)\sqrt{N_l}} \sum_{j=0}^{N_l-1} U^j \Big(\sum_{k\geq 1: N_{l-1} < n_k \leq N_l} e_k \Big)$$

weakly converge to a symmetrised Poisson distribution and for l even to the standard normal distribution. We can do so by defining, for l odd and $N_{l-1} < n_k \leq N_l$, $e_k = \pm r_k$ with probabilities $1/(2N_l)$ and $e_k = 0$ with probability $1 - 1/N_l$, where

$$\sum_{k\geq 1: N_{l-1} < n_k \leq N_l} r_k = b(N_l)\sqrt{N_l},$$

for l even we define e_k normally distributed with zero means and variances r_k^2,

$$\sum_{k\geq 1: N_{l-1} < n_k \leq N_l} r_k = b(N_l)\sqrt{N_l}.$$

By (14) and (15) we then get the convergence to the same laws of $1/(b(N_{2l})\sqrt{N_{2l}})S_{N_{2l}}(f)$ and $1/(b(N_{2l-1})\sqrt{N_{2l-1}})S_{N_{2l-1}}(f)$.

3. *Existence of* $(\Omega, \mathcal{A}, \mu, T)$.

We define, for $l = 1, 2, \dots$, $A_l = \{-1, 0, 1\}$ for l odd and $A_l = \mathbb{R}$ for l even, equipped with the probability measures ν_l such that $\nu_l(\{-1\}) = 1/(2N_l) = \nu_l(\{1\})$ and $\nu_l(\{0\}) = 1 - 1/N_l$ for l odd, $\nu_l = \mathcal{N}(0,1)$ for l even. For each l we define $\Omega_l = A_l^{\mathbb{Z}}$ equipped with the product measure $\mu_l = \nu_l^{\mathbb{Z}}$ and with the transformation T_l of the left shift. Then we put $\Omega = \bigtimes_{l=1}^{\infty} \Omega_l$, equipe it with the product measure $\mu = \bigotimes_{l=1}^{\infty} \mu_l$ and the product transformation T. The random variables e_k, $N_{l-1} < n_k \leq N_l$, will then be multiples of projections of Ω onto A_l.

This finishes the proof.

References

[B 1961] Billingsley, P., *The Lindeberg-Lévy theorem for martingales*, Proc. Amer. Math. Soc. **12** (1961), 788-792.

[D-L] Derriennic, Y and Lin, M., *The central limit theorem for Markov chains with normal transition operators, started at a point*, Probab. Theory Relat. Fields **119** (2001), 509-528.

[G] Gordin, M.I., *A central limit theorem for stationary processes*, Soviet Math. Dokl. **10** (1969), 1174-1176.

[G-Ho] Gordin, M.I. and Holzmann, H., *The central limit theorem for stationary Markov chains and various invariant splittings*, Stochastics and Dynamics **4** (2004), 15-30.

[G-L] Gordin, M.I. and Lifšic, B.A., *A remark about a Markov process with normal transition operator*, In: Third Vilnius Conference on Probability and Statistics **1** (1981), 147-148.

[Ha-He] Hall, P. and Heyde, C.C., *Martingale Limit Theory and its Application*, Academic Press, New York, 1980.

[I] Ibragimov, I.A., *A central limit theorem for a class of dependent random variables*, Theory Probab. Appl. **8** (1963), 83-89.

[K-V] Kipnis, C. and Varadhan, S.R.S., *Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions*, Comm. Math. Phys. **104** (1986), 1-19.

[Kl-Vo 1] Klicnarová, J. and Volný, D., *An invariance principle for non adapted processes*, preprint (2007).

[Kl-Vo 2] Klicnarová, J. and Volný, D., *Exactness of a Wu-Woodroofe's approximation with linear growth of variances*, preprint (2007).

[M-Wo] Maxwell, M. and Woodroofe, M., *Central limit theorems for additive functionals of Markov chains*, Ann. Probab. **28** (2000), 713-724.

[P-U] Peligrad, M. and Utev, S., *A new maximal inequality and invariance principle for stationary sequences*, Ann. Probab. **33** (2005), 798-815.

[V 2006] Volný, D., *Martingale approximation of non adapted stochastic processes with nonlinear growth of variance*, Dependence in Probability and Statistics Series: Lecture Notes in Statistics, Vol. 187 Bertail, Patrice; Doukhan, Paul; Soulier, Philippe (Eds.) (2006).

[Wo] Woodroofe, M., *A central limit theorem for functions of a Markov chain with applications to shifts*, Stoch. Proc. and their Appl. **41** (1992), 31-42.

[Wu-Wo] Wu, Wei Biao and Woodroofe, M., *Martingale approximation for sums of stationary processes*, Ann. Probab. **32** (2004), 1674-1690.

Département de Mathématiques, Université de Rouen, 76801 Saint Etienne du Rouvray, France

E-mail address: `Dalibor.Volny@univ-rouen.fr`